

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **04264985 A**

(43) Date of publication of application: **21.09.92**

(51) Int. Cl.

**G06F 15/62**  
**G06F 15/18**  
**G06F 15/62**  
**G06F 15/70**  
**H04N 7/18**

(21) Application number: **03026103**

(22) Date of filing: **20.02.91**

(71) Applicant: **HAMAMATSU PHOTONICS KK**

(72) Inventor: **HANASHIMA MASAOKI**  
**KURONO TAKEHIRO**

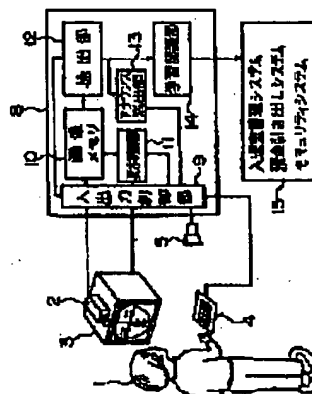
(54) **IMAGE RECOGNITION SYSTEM**

COPYRIGHT: (C)1992,JPO&Japio

(57) Abstract:

**PURPOSE:** To recognize a subject consisting of the whole and its constitutional part such as a face at high speed and with high accuracy by recognizing the subject by inputting the feature part of the subject to an individual neural network.

**CONSTITUTION:** The face of a person 1 to be inspected is image-picked up by directing a telecamera 2, and the image of the person 1 to be inspected is displayed on a display monitoring device 3. The person 1 to be inspected freezes the image and sends an instruction for input definition by a key operation from a keyboard input device 4, and also, announcing display whether or not the posture of the face of the person 1 to be inspected is correct is performed by, for example, a tone. Image data is extracted by a mask which extracts the whole and part of the subject, and it is supplied to the neural network at every extracted image data, thereby, learning and recognition are performed. Also, parallel processing is performed by constructing each neural network with different processors.



60649/00F



## Tele Connect Videosystemen

*Franciscusweg 277 - 1216 SG Hilversum - Nederland*

*Tel: 035 - 6282245      Fax: 035 - 6210885*

index

# SANYO

XP 2226856

## Frequently asked questions

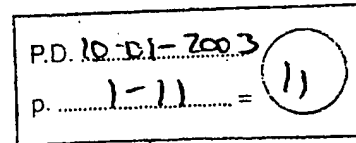
**Facial Templates:** XP-002226854

## 1. What is a facial template?

While one can always apply face recognition to two facial images, there are advantages to comparing instead a facial image against a facial template:

1. Speed of comparison.
2. Storage size.

Visionics offers templates optimized for speed and/or size, for both identification and verification.



**2. What are the types of facial templates used by Visionics for one to many searches (identification)?**

**Visionics currently uses two types of templates for identification:**

- a. A small template, currently 88 bytes, is used for fast searching over the entire query restricted sub-database. This is known as a vector template.
- b. A large template of size 3.5Kb, is used for an intensive search over the top 5% (or less) ordered matches during 1:n matching. This is known as a full template.

It is possible to use the vector templates alone in some cases.

**3. What are the types of facial templates used by Visionics for one to one matching (verification)?**

- cropped and scaled image of the person's face can be used as the template - known as a big canonical image. A compressed image of 100 to 300 bytes for verification.
- A highly compressed image of 100 to 300 bytes, a compressed face.

Using a big canonical image for verification avoids the processing time required to create a vector template. In addition, these images typically compress to 1.8K in size using JPEG compression with compression factor 80, and are smaller than the full template.

#### 4. Can the identity of the person be obtained from the vector or full templates?

No. The facial image cannot be reconstructed from these templates.

### 5. What is a canonical or big canonical image?



*The original image.*

From an original image, rotate, scale, and crop the image so that the eyes appear in the same position and the image size is 80 pixels by 100 pixels. This generates a canonical image shown below



*Canonical Image*

A big canonical image is similar except that more area around the face is kept so the image size is 100 x 125 pixels. This is a more robust template for verification due to internal algorithmic processes.



*Big Canonical Image*

**6. Why is Visionics face compression better than any other standard form of compression?**

It is designed specifically for faces. This allows for a much better compression of faces than generic compression algorithms. Furthermore, it is designed for use with Visionics face recognition technology.

**7. How small can facial image be compressed?**

Down to roughly 100 bytes.



From the original image, a big canonical image is created. The big canonical image is masked (the black rectangle with the elliptical hole) in preparation for compression. The compression is optimized to represent the inside of the ellipse.



*Uncompressed Masked Image*

Below is the masked image after compression



*1000 bytes*



*300 bytes*



*100 bytes*

The compression API also lets you view the full image, without the mask. Note that the non-facial portion does not reconstruct as well, but since it does not contribute to the person's identity from the Visionics algorithm point of view, the data is not important.



*100 bytes unmasked*

**8. How many templates or images should be stored for optimal performance for verification systems?**

For optimal performance during verification or surveillance, multiple varying images should be used to recognize a single person. For verification, three to nine images are typical for optimal performance.

---

## Face Recognition:

**1. What scientific method does FaceIt® use for face recognition?**

Fundamental to any face recognition system is the way in which faces are coded. FaceIt® uses Local Feature Analysis (LFA) to represent facial images in terms of local statistically derived building blocks.

LFA is a mathematical technique that is based on the realization that all facial images can be synthesized from an irreducible set of building elements. These elements are derived from a representative ensemble of faces using sophisticated statistical techniques.

They span multiple pixels (but are still local) and represent universal facial shapes, but are not exactly the commonly known facial features. In fact, there are many more facial building elements than there are facial parts. However, it turns out that synthesizing a given facial image, to a high degree of precision, requires only a small subset (12-40 characteristic elements) of the total available set. Identity is determined not only by which elements are characteristic, but also by the manner in which they are geometrically combined (i.e. their relative positions).

**2. How do changes in expression, such as smiling, frowning or blinking affect FaceIt® face recognition?**

LFA has advantages over earlier approaches. Compared with "eigenfaces", LFA face recognition is relatively insensitive with respect to changes in expression, including blinking, frowning, and smiling.

**3. Is FaceIt® face recognition sensitive to the growth of facial hair?**

No, LFA has enough redundancy and robustness to be able to compensate for mustache or beard growth.

**4. Is FaceIt® face recognition sensitive to hairstyle?**

No, the hair is not used as a local feature.

**5. Does FaceIt® use neural network technology?**

The algorithms have been "trained" on human faces to determine the correct significance of

each local feature. However, neural networks are not used in the usual sense.

**6. Can people be recognized at varying pose?**

For pose less than 10-15 degrees, there is no degradation in face recognition. From 15 to 35 degrees, the face recognition discrimination power decreases. Angled faces beyond 35 degrees do not match well to frontal faces with our current release technology. However, angled faces can be matched to other faces at the same angle as long as both eyes are clearly visible.

**7. Can people be recognized with eyeglasses on or off?**

Yes, the FaceIt® algorithm is able to compensate for eyeglasses.

**8. Is the technology sensitive to race or gender?**

No, special care has been taken to assure that Visionics face recognition technology is invariant with respect to race or gender.

**9. Can an image captured with a digital camera be matched against a digital video camera or scanner?**

Yes, internal algorithms are applied to make the matching relatively invariant with respect to input type.

**10. What are the major causes of face recognition failure?**

- a. Significant glare on eyeglasses that obstructs the eyes. Adjusting the lighting that causes the glare can typically solve this.
- b. Long hair obscuring the central part of the face.
- c. Poor lighting that would cause the face to be overexposed and low contrast.
- d. Lack of resolution, in pixels, of the face.

**11. Can the face recognition match accurately an image created with a photo-fit product (eg CD-Fit, comPHOTOfit, Suspect ID) against an actual image?**

If the photo-fit product produces smooth images, as opposed to high contrast sketch images, then our software works well. However, our software will not work as well with high contrast "cartoon-like" images.

**12. Can Visionics' face recognition match accurately an image created with an aging product against an actual image?**

- a. Our algorithms have been modified to better recognize infants and small children.
- b. We have studied the effect of aging from adolescence through adulthood using our technology and have found invariance with respect to aging beyond the completion of feature growth (roughly 13 years of age).
- c. If the aging product produces smooth images, as opposed to high contrast sketch images, then our software works well. However, our software does not work as well with high contrast "cartoon" images.

**13. How does one characterize the accuracy of face recognition?**

When performing recognition, one typically attempts to obtain a score above a fixed threshold. If the score is above or equal to the threshold, the person is recognized, otherwise the person is not recognized.

On the average, accuracy is characterized in terms of two probabilities at a given threshold.

1. False Acceptance Rate (FAR): The chance that an imposter will be recognized (obtain a higher score) at a certain threshold.
2. False Rejection Rate (FRR): The chance that the correct person will not obtain a score above a certain threshold.

Both the FAR and FRR are functions of threshold. The threshold where the two probabilities are the same is the Equal Error Rate (EER). For example, if the EER is 1%, that means 1% of the right people are rejected and 1% of the wrong people are accepted above a certain threshold.

**14. How are the FAR and FRR and Equal Error Rates determined?**

These numbers are determined by applying the face recognition algorithm to a large database of faces, where the correct matches have been pre-determined.

The accuracy of the face recognition system is strongly dependent on the database used for analysis. In the results below, we use the publicly available portion of the FERET (provided by the U.S. Army Research Lab, Washington D.C., USA) database. This database is available upon request.

**15. Do the FAR and FRR values depend upon the database used in the analysis?**

Yes. The False Rejection Rate (FRR) is strongly sensitive to the database used to calculate the result. This is because the rejection of an individual may occur simply because the image quality of the database is poor.

The False Acceptance Rate (FAR) is less sensitive to the image quality of the database. However, databases still vary in recognition "difficulty".

Therefore, both the FRR and FAR (to a lesser extent) are sensitive to the database used for analysis.

**16. Is there a standard database that can be used to judge all facial technologies on the same "playing field"?**

The closest to a standard database is the US Army FERET (Face Recognition Test) database. In the results below, we use the publicly available portion of the FERET database (provided by the U.S. Army Research Lab, Washington D.C., USA). This database is available upon request.

**17. What is the performance of Visionics face recognition on this database?**

The Equal Error Rate is .68% (less than one percent).

---

## Face Finding:

**1. What scientific method is used for face finding?**

FaceIt® face finding uses a combination of geometrical queues and pattern matching to find heads and facial features. Visionics face finding can detect simultaneously the presence of multiple faces in an image or in video frames, and can accurately determine the position of the faces.

**2. Does the user have to supply clues for the face finding to work?**

No, the entire face finding process is fully automated, continuous and functions in real-time on standard off-the-shelf processor. The user does not have to click on the image as a clue.

**3. Does the person have to be facing the camera for the face finding to work?**

No, the face finding technology can find a face as long as both eyes are clearly visible.

**4. Does the capturing program automatically measure the pose-offset angle of the face?**

No, the pose is not currently estimated.

**5. Can the distance from the face automatically be measured?**

Visionics technology returns the eye positions and subsequently the size of the face in the image.

**6. Does the face finding technology require the face to be of a certain size or in a certain position?**

No, Visionics technology finds faces at arbitrary scale (up to some resolution) and subsequently generates a re-scaled the face in a standard size for facial matching.

**7. How accurate is the face finding?**

The face finding technology finds faces and returns a score indicating the "goodness" of the face found (known as the alignment quality). This way, the small percentage (of order 1%) of aligned faces can be manually aligned to complete the template creation process on a large database. For example, to manually align a one million-person database at 10 alignments (selecting the eye positions) per minute and eight-hour days, it would take 208 man-days. On the other hand, it would only take 8 days to automatically align the faces at an alignment speed of one per second.

**8. Where is the face finding technology in the Visionics SDKs?**

In the ILocate COM object, the CVerify ActiveX control, and the CLocateXX C++ classes in the FaceIt® Library.

---

## Tracking:

**1. What is the scientific method used for tracking?**

Once the face is found, the tracking algorithm follows the face based on the person's geometrical facial characteristics and their flesh tone.

**2. How much can the face deviate from frontal pose before the tracking breaks?**  
The face and head can be tracked 360 degrees. That is, the person can turn completely around and the tracking will still follow the person.

**3. Are there any special hardware requirements for tracking?**  
The camera/digitizer system must be fast and, unlike basic face recognition, provide color information. A PCI capture system with high quality camera is recommended.

**4. Where is the tracking technology in the Visionics SDKs?**  
In the ITracking COM object, the CVerify ActiveX control, and the CLocateTR C++ class in the FaceIt Library.

---

## **Liveness:**

### **1. What is "liveness"?**

"Liveness" is the name given to solving the problem of misinterpreting a photo of a person as that person. Visionics provides two "Liveness" tests.

1. Single frame/image testing, which attempts to detect the characteristics of a photograph, such as rectangular borders.
2. Multiple frame/video testing, which is a challenge response system.

It asks the user to blink or smile, but not both at the same time .

### **2. How long does it take to test for "liveness"?**

Single frame "liveness" takes 1/10th of a second, multiple frame "liveness" takes user prompting and a few frames of video input, typically 2-3 seconds.

**3. Are there any special hardware requirements for "liveness"?**  
No.

**4. Where is the "liveness" test in the Visionics SDKs?**  
In the ILiveness COM object and the CLocateLV C++ class in the FaceIt library.

---

## **Image Quality:**

### **1. What is Image Quality?**

Image Quality is a test given to a facial image after capture but before the image is permanently stored in a database, to verify that the facial image will be useful for face recognition. This test gives the administrator a chance to take another picture if the quality test fails.

**2. How long does it take to test for Image Quality?**  
Less than 1/10th of a second.

**3. What are the Image Quality tests?**

1. Quality of the face found in the image.
  2. Size and position of face. Reject faces if cut off the image or if too small.
  3. Overexposure. Verify that the facial image is not overexposed (too dark).
  4. Underexposure: Verify that the facial image is not too light.
- 

## **Memory and Speed of Face Finding and Recognition:**

### **1. How fast is the face recognition search speed?**

Reference computer: 500 MHz Pentium III CPU. Using vector templates, the raw search speed is 15 million matches per minute average, 47 million peak optimized.

Using the full templates, the raw search speed is 15,000 matches per minute.

The search speed is roughly linear with CPU speed. For example, with a 733 MHz CPU, the peak search speed would be 68 million per minute, better than 1 million per second.

### **2. What is the face recognition speed for a hierarchical search, where vector templates are searched on the bulk of the database and full templates are?**

The formula for a hierarchical search using first the small templates and subsequently the larger templates is:

$N = N_v / (1 + F * N_v / N_i)$ , where

$N_v$  = the number of searches per minute in vector mode,

$N_i$  = the number of searches per minute in intensive mode,

$F$  = fraction of database searched in intensive mode.

For example,  $N_v=45000K$ ,  $N_i=15K$ ,  $f = 0.005$  (0.5%) yields an overall search speed of 2,800,000 searches per minute.

### 3. How much memory is required to execute the face recognition algorithm?

1. For matching two vectors, less than 2kilobytes of RAM are required.
2. For matching an image to a full template, roughly 1.4 megabytes of RAM is required.

### 4. How long does it take to create a template?

On a 500 MHz Pentium III CPU, the template creation times are:

1. Vector Creation time: 1 second.
2. Full template creation time: 1/30th of a second.
3. (Big) Canonical Image creation time: 1/50th of a second.

### 5. How much RAM is required to create a template?

1. For vector creation, roughly 2 megabytes of RAM are required.
2. For full template creation, roughly 0.5 megabytes of RAM is required

### 6. How fast is the face finding search speed?

Face finding on a typical input image of 400x300 pixels takes between 0.3 and 1.0 seconds, depending upon the size and quality of the face in the image.

Note: This time can be reduced if the rough sizes of the faces in the images are known beforehand. Then the face finding algorithm does not need to spend time looking for faces of all sizes.

### 7. How much memory is required to execute the face finding algorithm?

Roughly 5 megabytes of RAM are required to analyze typical image for a face.

---

## FaceIt DB/Sentinel/Surveillance:

### 1. What is the search speed of the FaceIt Search Engine in FaceIt DB?

The maximum search speed is roughly 15 million per minute per CPU (500 MHz).. The speed is less than the maximum quoted speed of 47 million due to extra checks performed to maintain data integrity, and due to the usage of Microsoft Access and the image storage database.

### 2. What databases are used by FaceIt DB/Sentinel/Surveillance?

FaceIt DB/Sentinel/Surveillance uses two databases systems.

1. A proprietary flat file system to store and quickly retrieve facial templates.
2. Microsoft Access to store images and personal information.

Microsoft Access can slow down dramatically during data insertions and deletions when the size of the image database grows significantly beyond 30 thousand. Since FaceIt DB/Sentinel/Surveillance connects to Access via the Microsoft ODBC layer, connecting to other ODBC compatible databases, such as Microsoft SQL server can solve this problem. However, non-Access databases require a custom setup to work with DB/Sentinel/Surveillance.

---

## Database Storage:

### 1. Do you have any specific storage requirements in regards to 1:N matching?

Yes, for efficient 1:N matching, a facial template must be created and stored.

Facial template creation involves finding the face in the image, and then processing the found face into a byte array that can be later used for efficient facial searching.



Therefore, during the photo storage process, CPU resources need to be allocated to perform the face finding, and subsequently storage resources must be allocated to store the facial search template.

## **2. What does the template creation process entail?**

Visionics face finding technology can automatically find a human face in a single image or in a live video feed. The face finding process occurs in real-time (roughly 1 second).

This allows for the automatic creation of facial templates in large existing databases without human intervention.

After face finding, template creation can be executed on the found face, or template creation can be executed separately on a manually found face.

## **3. Do watermarked images affect the performance of the face recognition?**

No. The correct concept is to think of Visionics as a technology provider independent of a specific platform or a specific database format, and FaceIt® technology as a filter system that converts facial images into facial templates and the combination of facial images and search templates into ordered lists.

Visionics offers a FaceIt DB Enterprise custom solution that incorporates a proprietary database system for template storage. This product combines the technology with database storage and management on the Windows NT platform. However, the technology can be used directly through a Visionics developer kit API, with no specification of database. Basic Visionics technology does not include a database.

## **4. Can your technology work with RDBMS, ODBMS, and ORDBMS?**

RDBMS Examples: Oracle, MS SQL Server, Access

ODBMS Examples: Jasmine, Poet, Objectivity, ObjectStore

ORDBMS Examples: Informix

The systems listed above include interfaces for imbedded SQL processing that would interface well with our technology.

For example, we offer an Informix Datablade that can perform the face finding, template creation and face recognition tasks internal to the Informix system.

The two possible solutions for processing during face finding, template creation, and face recognition are:

a. Process internal to the database using imbedded SQL processing via a Datablade-like plug-in.

b. Store only pointers to template in the image databases and store/process the template in a custom external database system. We offer the FaceIt DB Enterprise custom face recognition/database solution.

The choice of system depends upon the existing infrastructure.

If the image storage systems are multi-vendor (for example, using both Oracle and Microsoft or both Windows NT and Unix platforms), b. may be the most cost-effective centralized solution.

For single vendor database single vendor platform infrastructures, a. may be the most cost effective solution.

## **5. Can you discuss further an implementation strategy for a large-scale database system using face recognition?**

We recommend a multi-tier solution:

Tier1: Client web-based front end. The clients submit images and perform queries including face recognition through their browsers.

Tier2: Web and Application Server. Software provides database connectivity - converts custom tags in the web pages into database SQL instructions.

a. Netscape Enterprise Server 3.0. Developers have provided native database connectivity to Oracle, Sybase, Informix, DB2, and connectivity to other data sources through ODBC.

b. Microsoft Internet Information Server: Designed to provide Connectivity to MS SQL and Microsoft Message Queue Server layers via Active Server Pages (ASP) technology.

In addition this tier connects via an internal network messaging system, such as Microsoft Message Queue Server (MMQS) or IBM MQSeries, to the image server and the face recognition server.

Tier3: Image and Person Info Server Stores digitized images and tags denoting gender, citizenship, criminal status (if any), etc - all personal information which might be used to constrain a face recognition search or image lookup. Communicates to Tier2 via the network-messaging layer.

Tier4: Face Recognition and Face Finding Server: Performs all template creation (probably using automated face finding) and all face recognition queries. Communicates to Tier 3 and 2 via the network-messaging layer.

In the context of Tier 4, there are two solutions currently implemented by Visionics:

a. Process internal to the database using imbedded SQL processing via a "Datablade" like plug-in. The database software handles I/O and queuing. Visionics would provide the plug-in.

b. Process in the context of a layer such as MMQS, with queuing and I/O handled by the fully scalable Visionics FaceIt DB Enterprise solution.

In both cases, the same fundamental technology would be used .

#### **6. Volume: Is there a limit to the volume for 1:N matching?**

No.

#### **7. What are the system requirements for a large-scale search system?**

In the context of an external facial system with a legacy database system storing images, a set of workstations for facial queries, and a backend system for facial processing, we calculate below the number of computers required to meet a specified search specification:

Computer: 500 MHz CPU system with base RAM of 512MB, base and base disk space of 10GB.

One can replace a stand-alone computer with a processor in a dedicated multiple processing system if the system includes the RAM and disk as specified above to each CPU.

Disk: This disk storage requirement of roughly 4K per overall facial templates per image will allow storage of 2.8 million individuals.

RAM: The RAM storage requirement of less than 128 bytes per person translates into 320MB of RAM for fast template storage.

CPU: The overall search speeds are roughly 2.8-million/per minute/per CPU. (see the performance section below). This translates into a formula for the number of computers:

$N_{\text{computers}} =$

$(\text{Total Population size}) / (2.8 \text{ million} * \text{minutes/search}), \text{ where minutes/search} \leq 1.$

For example, for searching through 20 million in one minute, the number of computers is 8.

In addition to search and alignment engines, one addition computer may be required to act as a master controller.

---

## **Image Input:**

### **1. Do you have any recommendations for digital cameras to use for the generation of a database of images for facial searching?**

Most personal computer based video capture devices are inadequate for high-quality facial recognition results. We recommend instead the use of a "MegaPixel" digital camera with "flash on demand", or a high quality (300 DPI or above) scan of a good quality photograph.

A good example is the KODAK DC260 Zoom Digital Camera, 1536 x 1024 resolution, 3X optical plus 2X digital zoom lens. It is capable of flash sync and manual exposure adjustment. The resolution is less important than the flash sync and the quality of the exposure.

In general, there are a large number of adequate digital cameras for the task of photo imaging for face recognition tasks or identification card creation. The correct solution for a given application depends upon the price and availability at the time of the contract award. The price for such cameras has dropped by a factor of two roughly every two years and the resolution has also increased by a factor of four over the past year for low-end systems. We expect the price trend to continue, while the resolution will remain roughly in the mega-pixel range over the next two years .

**2. What sort of image enhancements might be required based on the photo input?**  
The internal algorithms in the Visionics FaceIt® technology have image enhancement technology. No external image enhancement is required.

**3. What is the full definition of best quality photo input?**  
The recommended image format is JPEG, with 24-bit color or 256 shades of gray color depths. The JPEG compression for the photos should be at maximum roughly 15:1 ("quality" 80 for JPEG).

The best images for processing are 300 x 400 x 24-bit resolution. The entire head, including hair and part of the neck should constitute the bulk area of the image, and there should be roughly 100 pixels of data from eye to eye.

The best faces for matching are directly facing forward under controlled, balanced lighting, with eyes open and a neutral expression like a mug shot photograph.

Under no conditions should the head/neck be cut off in any way, as this will compromise the quality of the automatic face finder.

Other situations to be avoided include glare on eyeglasses that obscure the eyes, sunglasses, closed eyes, mouths open during speech, strong smiles with exposed teeth, and variations in pose (left/right) or tilt (up/down) beyond 10 degrees in any direction.

Larger resolution photos can be accepted but generally do not increase accuracy to any measurable extent and take longer to process.

**4. What is the minimum photo input specifications required for the Visionics technology maintain effectiveness?**  
The minimum quality image would be 100 x 125 x 256 Grayscale JPEG with 15:1 compression, and 20 pixels from eye to eye.

**5. If the source and target image formats are different, will the results be the same as matching using the same formats?**  
Visionics FaceIt® technology uses the grayscale component of the input images. There will be no difference between use of color or grayscale formats as long as the grayscale color band is robust to 8 bits in both cases.

---

## Video Input:

**1. What video device standards do you support on the Windows platform?**  
We currently support in our products the Microsoft Video For Windows (VFW) standard. Nearly all video input devices designed for the windows platform support VFW. [Using our low-level API calls provided by our SDKs however, a programmer can connect any video device to our algorithms.]

**2. Do you support USB cameras?**  
If the camera manufacturer provides a VFW software driver for their device, then that device will work with our products and developer software. Our software does not connect directly to hardware.

**3. What capture hardware and cameras do you recommend for use with your technology?**  
PCI capture cards  
The Integral Technologies FlashBus MVLite PCI capture card  
The Winnov PCI capture card: [www.winnov.com](http://www.winnov.com) works reasonably well with the Winnov camera, and takes standard video input as well.  
The 3COM big picture PCI card and camera is a great low cost system

USB Desktop Cameras  
The Kodak DVC323  
The Winnov USB camera

Parallel Port Cameras  
The Vicam camera

Desktop Cameras with Video Output  
For a large list, see the NetMeeting web site

The Sanyo CCD desktop camera is a good example of a high quality/low cost desktop camera

We also like the Howard HA6800 with Zoom lens for high quality video input for desktop systems

For demos, teleconferencing, or high-end use, we like the more expensive Sony EVI-D30 pan tilt zoom auto-focus camera

The Connectix Quickcam series of cameras is not recommended for use with face recognition due to below average performance.

#### Fixed Field Surveillance Cameras

#### Pan-Tilt-Zoom Surveillance Cameras

We have no recommendations for pan/tilt/zoom surveillance camera systems at this time.

#### **4. What sort of image enhancements might be required based on the video input?**

You should enable automatic gain control on the video camera.

#### **5. How can the quality of video be controlled to ensure optimal results?**

Avoid including a bright light source in the video field of view such as the sun, or when indoors, a window in the background field of view.

In general, avoid situations that will generate a photographic back lighting problem .

#### **6. What are the recommended video digitizer settings?**

Analog video input must be digitized (resolved into pixels) before it can be processed for face finding and face recognition. The default digitizer setting for most desktop systems is 160x120 pixels.

This resolution is not recommended for use with Visionics technology. Instead we recommend:

For desktop verification: RGB888 320x240 pixels.

For video surveillance: RGB888 640x480 pixels.

The setting can be adjusted via the manufacturer's digitizer options dialog box, which is a standard component of VFW software drivers .

#### **7. What is the full definition of best quality video input for face finding and face recognition?**

The video digitizer settings are RGB888 640x480.

The face and neck are clearly visible in the video field, and roughly the face takes up one-third of the video field. Under no conditions should the head be cut off in any way, as this will compromise the abilities of the automatic face finder.

The face should not be so close to the camera as to cause a "fish-eye" effect where the nose is larger than normal.

The best faces for matching are directly facing forward under controlled, balanced lighting, with eyes open and a neutral expression like a mug shot photograph.

Other situations to be avoided include glare on eyeglasses that obscure the eyes, sunglasses, closed eyes, mouths open during speech, strong smiles with exposed teeth, and variations in pose (left/right) or tilt (up/down) beyond 15 degrees in any direction .

#### **8. What are the minimum video input specifications required for the Visionics technology to maintain good effectiveness?**

There should be minimum 20 pixels between the eyes. The person's pose should be within 35 degrees of frontal, within 15 if possible.

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 03/11429

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06K9/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	"Frequently asked questions" TELE CONNECT VIDEOSYSTEMEN, XP002226854 Hilversum, Nederland	1-6, 13
Y	Aus dem Internet am 10. Januar 2003 von <a href="http://www.teleconnect.nl/faces-pages/face-s-techbes-als-face-it-faq.pdf">http://www.teleconnect.nl/faces-pages/face-s-techbes-als-face-it-faq.pdf</a> ; gemäß Dokumentenquelle erstellt und zuletzt modifiziert am 23. Juli 2002 page 1 -page 6	7-12, 14-18
Y	US 2001/046330 A1 (SHAFFER STEPHEN L ET AL) 29 November 2001 (2001-11-29) page 4, paragraph 30 -page 5, paragraph 33	7-12, 14-18
A	EP 0 984 386 A (SHARP KK) 8 March 2000 (2000-03-08) paragraph '0096! - paragraph '0106!	1-18
	-/-	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- 'A' document defining the general state of the art which is not considered to be of particular relevance
- 'E' earlier document but published on or after the international filing date
- 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- 'O' document referring to an oral disclosure, use, exhibition or other means
- 'P' document published prior to the international filing date but later than the priority date claimed

- 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- 'G' document member of the same patent family

Date of the actual completion of the international search

16 January 2004

Date of mailing of the international search report

06/02/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentean 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 81 851 epo nl  
Fax: (+31-70) 340-3016

Authorized officer

Müller, M

# INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 03/11429

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 963 656 A (YAO YI-SHENG ET AL) 5 October 1999 (1999-10-05) abstract; claim 1; figure 8 column 2, line 31 - line 42 column 9, line 5 - line 57	1-18
A	WO 02 35453 A (LEE DEUNGGU ;ALPHA ENGINEERING CO LTD (KR)) 2 May 2002 (2002-05-02) abstract; figures 1,9 page 4, line 35 -page 6, line 21	1-18
A	PATENT ABSTRACTS OF JAPAN vol. 017, no. 051 (P-1479), 2 February 1993 (1993-02-02) & JP 04 264985 A (HAMAMATSU PHOTONICS KK), 21 September 1992 (1992-09-21) the whole document	1-18

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/EP 03/11429

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2001046330	A1	29-11-2001	NONE
EP 0984386	A	08-03-2000	GB 2341231 A 08-03-2000 EP 0984386 A2 08-03-2000 JP 2000082147 A 21-03-2000 US 6633655 B1 14-10-2003
US 5963656	A	05-10-1999	NONE
WO 0235453	A	02-05-2002	KR 2001000861 A 05-01-2001 AU 1104902 A 06-05-2002 WO 0235453 A1 02-05-2002
JP 04264985	A	21-09-1992	NONE